



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ**  
**INSTITUTO DE BIODIVERSIDADE E FLORESTAS**



Estatística Experimental

# Análise de Variância

Prof. Rafael Rode

# ASSUNTOS

- O que é a Análise de Variância
- Delineamento Inteiramente Casualizado
- Suposições da ANOVA
  - Normalidade dos dados
  - Homogeneidade das variâncias

# ANÁLISE DE VARIÂNCIA

## PROBLEMA:

Avaliar se dois ou mais tratamentos diferem estatisticamente

MÉTODO ESTATÍSTICO



ANÁLISE DE VARIÂNCIA

ANOVA

# ANÁLISE DE VARIÂNCIA

## ➤ Hipóteses da ANOVA

Quando se efetua uma análise de variância, a seguinte hipótese está sendo testada:

$H_0$ : os tratamentos têm os mesmos efeitos,  
ou seja,  $t_1 = t_2 = \dots = t_k$ .

$H_1$ : pelo menos dois tratamentos têm efeitos diferentes.

# ANÁLISE DE VARIÂNCIA

A ANOVA foi introduzida  
*Sir Ronald Aylmer Fischer*  
quando trabalhou na Estação  
Experimental de Rothamsted  
de 1919 a 1933.



Nacionalidade	 Inglês
Nascimento	17 de fevereiro de 1890
Local	Londres
Morte	29 de julho de 1962 (72 anos)
Local	Adelaide
	<b>Atividade</b>
Campo(s)	Genética evolutiva, estatística
Prêmio(s)	Medalha Real (1938), Medalha Guy de Ouro (1946), Medalha Copley (1955), Medalha Darwin-Wallace (1958)

# ANÁLISE DE VARIÂNCIA

- Em uma série de observações é verificado a Variação entre os dados coletados.
- Esta **Variação Total** pode ser atribuída à:
  - **Causas conhecidas**
  - **Causas Desconhecidas – ou Não Controláveis**
- As Causas Desconhecidas são isoladas na ANOVA e recebem o nome de **Erro** ou **Resíduo**.

# ANÁLISE DE VARIÂNCIA

- Então, a variação que contribui para o erro experimental pode ser de dois tipos:
  - a) Inerente à própria variabilidade do material experimental;
  - b) Proveniente da falta de uniformidade do ambiente que é conduzido o experimento.

# ANÁLISE DE VARIÂNCIA

## ➤ FONTE OU ORIGEM DAS VARIAÇÕES

VARIAÇÃO TOTAL

```
graph TD; A[VARIAÇÃO TOTAL] --> B[Fonte de Variação]; B --> C[CONHECIDA]; B --> D[DESCONHECIDA]; C --- E[Representam a variação ENTRE amostragens (tratamentos)]; D --- F[Representam a variação DENTRO de amostragens (erro ou resíduo)];
```

Fonte de Variação

CONHECIDA

Representam a variação  
**ENTRE** amostragens  
(tratamentos)

DESCONHECIDA

Representam a variação  
**DENTRO** de amostragens  
(erro ou resíduo)

# DELINEAMENTO INTEIRAMENTE CASUALISADO

## ➤ MODELO ESTATÍSTICO BÁSICO

Assim, uma observação pode ser representada por um modelo estatístico **aditivo** da média mais os tratamentos e mais os resíduos (erro):

$$y_{ij} = \mu + t_i + e_{ij}$$

Em que:

$y_{ij}$  = valor observado da parcela que recebeu o tratamento  $i$  na repetição  $j$ ;

$\mu$  = média da população;

$t_i$  = efeito do tratamento  $i$  aplicado na parcela;

$e_{ij}$  = erro ou efeito dos fatores não-controlados na parcela.

# ANÁLISE DE VARIÂNCIA - DIC

- Cada resultado é então colocado em uma tabela, chamada de *Tabela de Análise de Variância*:

## Tabela de Análise de Variância para DIC

Fonte de Variação (F.V.)	GL	SQ	QM	F
Tratamentos <i>(entre amostragens)</i>	$i-1$	SQTr	$SQTr / (i-1)$	$QMTr / QMR$
Resíduos <i>(dentro de amostragens)</i>	$(n-1)-(i-1)$	$SQT - SQTr$	$SQR / (j-1)$	
<b>TOTAL</b>	<b><math>n-1</math></b>	<b>SQT</b>		

# ANÁLISE DE VARIÂNCIA - DIC

SOMA DE QUADRADOS TOTAL:

$$SQT = \sum_{i=1}^i \sum_{j=1}^j y_{ij}^2 - C$$

**C = correção**

$$C = \left( \sum_{i=1}^i \sum_{j=1}^j y_{ij} \right)^2 / n$$

# ANÁLISE DE VARIÂNCIA - DIC

SOMA DE QUADRADOS DOS TRATAMENTOS:

$$SQTr = \sum_{i=1}^i \frac{T_i^2}{j} - C$$

$$T_i^2 = \left( \sum_{j=1}^j y_{ij} \right)^2$$

SOMA DE QUADRADOS DOS RESÍDUOS:

$$SQR = SQT - SQTr$$

# ANÁLISE DE VARIÂNCIA - DIC

QUADRADO MÉDIO DOS TRATAMENTOS:

$$QMTr = \frac{SQTr}{GLTr}$$

QUADRADO MÉDIO DOS RESÍDUOS:

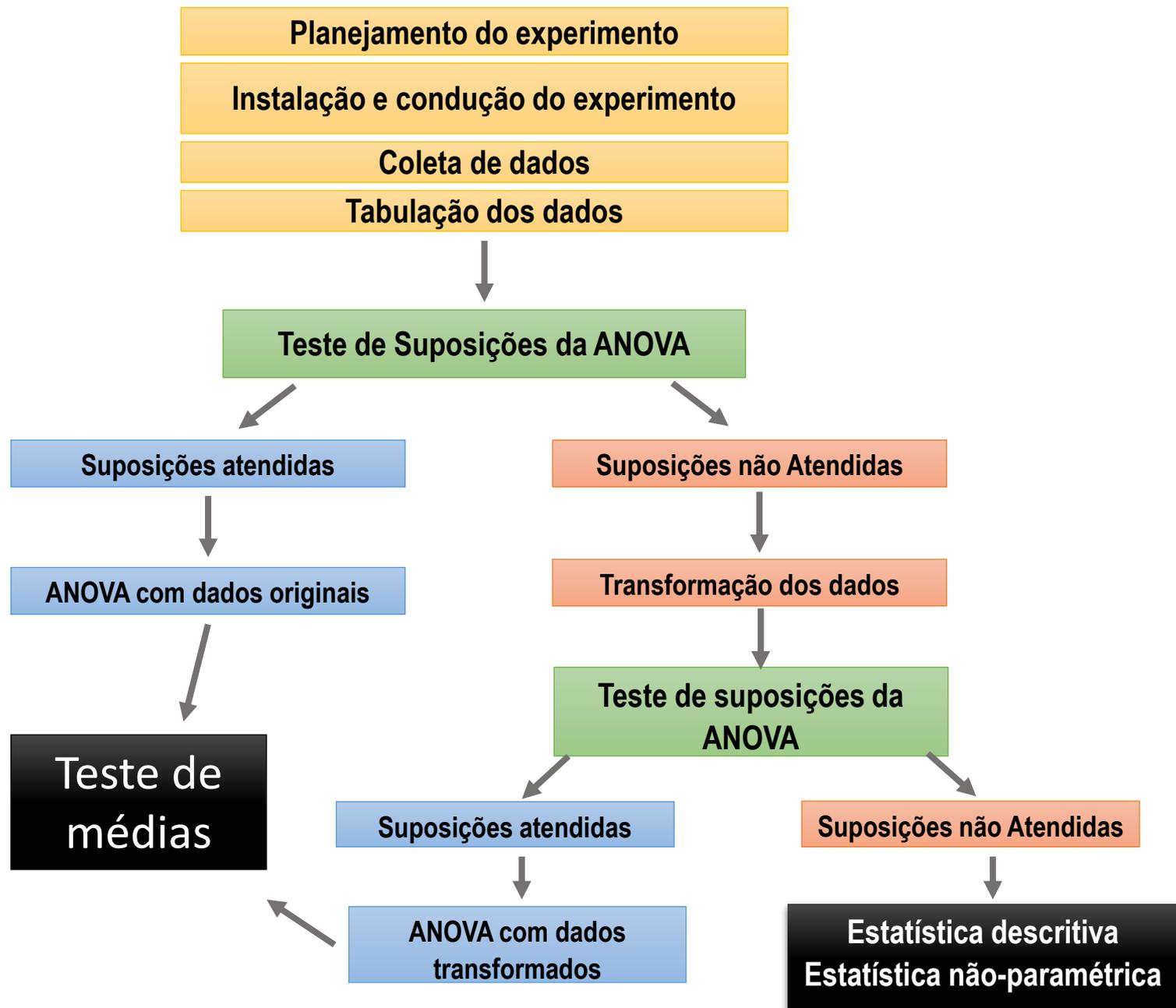
$$QMR = \frac{SQR}{GLR}$$

# ANÁLISE DE VARIÂNCIA - DIC

VALOR DE F:

$$F = \frac{QMr}{QMR}$$

# Fluxograma para experimentação e análise univariada



# SUPOSIÇÕES DA ANOVA

## ➤ O que é?

- São pressuposições ou premissas para realizar a análise de variância.

*O pesquisador deve verificar se os dados coletados atendem estas suposições antes de aplicar o teste F.*

## ➤ Para quê?

- Para que a Análise de Variância tenha validade;
- Para evitar conclusões sem justificativas;
- Para permitir conclusões importantes.

# SUPOSIÇÕES DA ANOVA

- 1. Os efeitos dos fatores que ocorrem no modelo devem ser aditivos (aditividade).**

$$y_{ij} = \mu + t_i + e_{ij}$$

Aditividade tem grande importância, mas em geral, ela já se verifica nos delineamentos experimentais

# SUPOSIÇÕES DA ANOVA

## 2. Os erros $e_{ij}$ devem ser independentes (independência).

Cada observação possui um erro que deve ser independente dos demais.

O princípio da casualização assegura a validade da estimativa do erro experimental, pois permite uma distribuição independente do mesmo.

# SUPOSIÇÕES DA ANOVA

## 3. Os erros $e_{ij}$ devem seguir a distribuição normal (normalidade).

- Quando: os erros ou desvios (devido a fatores não controlados) não seguem uma distribuição normal, os dados também não o são.
- Testes : Qui-quadrado, Lilliefors, Anderson-Darling, Kolmogorov-Smirnov, Shapiro-Wilk.
- Para corrigir a normalidade, a variável de estudo pode ser transformada matematicamente.

# SUPOSIÇÕES DA ANOVA

**4. Os erros  $e_{ij}$  devem ter variância comum (homogeneidade de variâncias ou homocedasticidade).**

- A variância de um tratamento deve ser semelhante aos demais tratamentos, ou seja, as variâncias dos tratamentos devem ser homogêneas.

Testes: Bartlett, Levene

# SUPOSIÇÕES DA ANOVA

Resumindo:

Para que uma análise de variância possa ser efetuada (e ser válida), é necessário que:

1. os dados possuem Distribuição Normal;
2. a variância dos tratamentos sejam homogêneas.

# Normalidade dos Erros

Se os dados a serem analisados tem distribuição normal, significa que o erros também tem.

## Testes de Normalidade (KS, Lilliefors, Anderson D. etc)

*Testes de normalidade são usados para determinar se um conjunto de dados de uma dada variável aleatória, é bem modelada por uma distribuição normal.*

### Hipóteses:

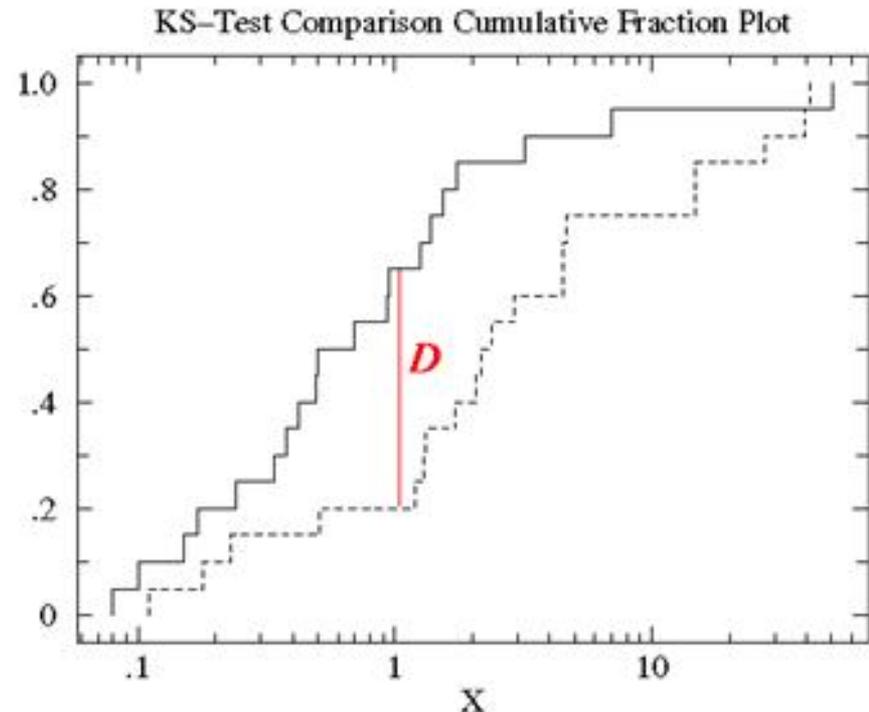
- $H_0$ : Os dados seguem uma distribuição normal
- $H_1$ : Os dados não seguem uma distribuição normal

# Normalidade dos Erros

## *Teste de Kolmogorov-Smirnov*

Este teste observa a máxima diferença entre a distribuição normal assumida para os dados e a distribuição empírica dos dados.

$$D_n = \sup_x |F(x) - F_n(x)|$$



## Valores Críticos da Distribuição da Estatística $D_n$ (Kolmogorov-Smirnov)

Os valores tabelados correspondem aos pontos  $D_{n,\alpha}$  tais que:  $P(D_n \geq D_{n,\alpha}) = \alpha$ .

n	$\alpha$				
	0.20	0.10	0.05	0.02	0.01
1	0.900	0.95	0.975	0.990	0.995
2	0.684	0.776	0.842	0.900	0.929
3	0.565	0.636	0.708	0.785	0.829
4	0.493	0.565	0.624	0.689	0.734
5	0.447	0.509	0.563	0.627	0.669
6	0.410	0.468	0.519	0.577	0.617
7	0.381	0.436	0.483	0.538	0.576
8	0.358	0.410	0.454	0.407	0.542
9	0.339	0.387	0.430	0.480	0.513
10	0.323	0.369	0.409	0.457	0.489
11	0.308	0.352	0.391	0.437	0.468
12	0.296	0.338	0.375	0.419	0.449
13	0.285	0.325	0.361	0.404	0.432
14	0.275	0.314	0.349	0.390	0.418
15	0.266	0.304	0.338	0.377	0.404
16	0.258	0.295	0.327	0.366	0.392
17	0.250	0.286	0.318	0.355	0.381
18	0.244	0.279	0.309	0.346	0.371
19	0.237	0.271	0.301	0.337	0.361
20	0.232	0.265	0.294	0.329	0.352

n	$\alpha$				
	0.20	0.10	0.05	0.02	0.01
21	0.226	0.259	0.287	0.321	0.344
22	0.221	0.253	0.281	0.314	0.337
23	0.216	0.247	0.275	0.307	0.330
24	0.212	0.242	0.269	0.301	0.323
25	0.208	0.238	0.264	0.295	0.317
26	0.204	0.233	0.259	0.290	0.311
27	0.200	0.229	0.254	0.284	0.305
28	0.197	0.225	0.250	0.279	0.300
29	0.193	0.221	0.246	0.275	0.295
30	0.190	0.218	0.242	0.270	0.290
31	0.187	0.214	0.238	0.266	0.285
32	0.184	0.211	0.234	0.262	0.181
33	0.182	0.208	0.231	0.258	0.277
34	0.179	0.205	0.227	0.254	0.273
35	0.177	0.202	0.224	0.251	0.269
36	0.174	0.199	0.221	0.247	0.265
37	0.172	0.196	0.218	0.244	0.262
38	0.170	0.194	0.215	0.241	0.258
39	0.168	0.191	0.213	0.238	0.255
40	0.165	0.189	0.210	0.235	0.252

# Normalidade dos Erros

## *Teste de Kolmogorov-Smirnov*

Conclusão:

Se  $D_n > D_{tab}$  (D crítico), rejeitamos a hipótese da nulidade dos dados ( $H_1$ ).

Caso contrário, não rejeitamos a hipótese de Normalidade ( $H_0$ ).

# Homogeneidade das variâncias

## ***Teste de variâncias (Bartlett, Levene etc)***

Use um teste para variâncias iguais para testar a igualdade de variâncias entre populações ou níveis de fatores.

Muitos procedimentos estatísticos, como a análise de variância (ANOVA) e regressão, assumem que embora amostras diferentes possam vir de populações com médias diferentes, elas possuem a mesma variância.

# Homogeneidade das variâncias

## *Teste de variâncias (Bartlett, Levene etc)*

Hipóteses:

- $H_0$ : As variâncias dos fatores são iguais
- $H_1$ : Nem todos os fatores tem variâncias iguais

# Homogeneidade das variâncias

## *Teste de Bartlett*

Em um experimento é testado a igualdade das variâncias dos tratamentos (homogeneidade).

Para o teste os dados devem apresentar distribuição normal.

Conclusão:

Se  $B_{calc} > B_{tab}$  (B crítico), rejeitamos a hipótese de variâncias iguais dos tratamentos (rejeita  $H_0$ )

Caso contrário, não rejeitamos a hipótese de homogeneidade das variâncias ( $H_0$ ).

**Tabela 5 - Distribuição de qui-quadrado:  $\chi^2(n)$ . Valores críticos de qui-quadrado tais que  $P(\chi^2 > \chi_c^2) = P$**

n	p=99%	98%	97,5%	95%	90%	80%	70%	50%	30%	20%	10%	5%	4%	2,5%	2%	1%
1	0,0 <sup>3</sup> 16	0,0 <sup>3</sup> 63	0,001	0,004	0,016	0,064	0,148	0,455	1,074	1,642	2,706	3,841	4,218	5,024	5,412	6,635
2	0,020	0,040	0,051	0,103	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	6,438	7,378	7,824	9,210
3	0,115	0,185	0,216	0,352	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	8,311	9,348	9,837	11,345
4	0,297	0,429	0,484	0,711	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	10,026	11,143	11,668	13,277
5	0,554	0,752	0,831	1,145	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,070	11,644	12,832	13,388	15,086
6	0,872	1,134	1,237	1,635	2,204	3,070	3,828	5,348	7,231	8,558	10,645	12,592	13,198	14,449	15,033	16,812
7	1,239	1,564	1,690	2,167	2,833	3,822	4,671	6,346	8,383	9,803	12,017	14,067	14,703	16,013	16,622	18,475
8	1,646	2,032	2,180	2,733	3,490	4,594	5,527	7,344	9,524	11,030	13,362	15,507	16,171	17,534	18,168	20,090
9	2,088	2,532	2,700	3,325	4,168	5,380	6,393	8,343	10,656	12,242	14,684	16,919	17,608	19,023	19,679	21,666
10	2,558	3,059	3,247	3,940	4,865	6,179	7,267	9,342	11,781	13,442	15,987	18,307	19,021	20,483	21,161	23,209
11	3,053	3,609	3,816	4,575	5,578	6,989	8,148	10,341	12,899	14,631	17,275	19,675	20,412	21,920	22,618	24,725
12	3,571	4,178	4,404	5,226	6,304	7,807	9,034	11,340	14,011	15,812	18,549	21,026	21,785	23,337	24,054	26,217
13	4,107	4,765	5,009	5,892	7,042	8,634	9,926	12,340	15,119	16,985	19,812	22,362	23,142	24,736	25,472	27,688
14	4,660	5,368	5,629	6,571	7,790	9,467	10,821	13,339	16,222	18,151	21,064	23,685	24,485	26,119	26,873	29,141
15	5,229	5,985	6,262	7,261	8,547	10,307	11,721	14,339	17,322	19,311	22,307	24,996	25,816	27,488	28,259	30,578
16	5,812	6,614	6,908	7,962	9,312	11,152	12,624	15,338	18,418	20,465	23,542	26,296	27,136	28,845	29,633	32,000
17	6,408	7,255	7,564	8,672	10,085	12,002	13,531	16,338	19,511	21,615	24,769	27,587	28,445	30,191	30,995	33,409
18	7,015	7,906	8,231	9,390	10,865	12,857	14,440	17,338	20,601	22,760	25,989	28,869	29,745	31,526	32,346	34,805
19	7,633	8,567	8,906	10,117	11,651	13,716	15,352	18,338	21,689	23,900	27,204	30,144	31,037	32,852	33,687	36,191
20	8,260	9,237	9,591	10,851	12,443	14,578	16,266	19,337	22,775	25,038	28,412	31,410	32,321	34,170	35,020	37,566
21	8,897	9,915	10,283	11,591	13,240	15,445	17,182	20,337	23,858	26,171	29,615	32,671	33,597	35,479	36,343	38,932
22	9,542	10,600	10,982	12,338	14,041	16,314	18,101	21,337	24,939	27,301	30,813	33,924	34,867	36,781	37,659	40,289
23	10,196	11,293	11,688	13,091	14,848	17,187	19,021	22,337	26,018	28,429	32,007	35,172	36,131	38,076	38,968	41,638
24	10,856	11,992	12,401	13,848	15,659	18,062	19,943	23,337	27,096	29,553	33,196	36,415	37,389	39,364	40,270	42,980
25	11,524	12,697	13,120	14,611	16,473	18,940	20,867	24,337	28,172	30,675	34,382	37,652	38,642	40,646	41,566	44,314
26	12,198	13,409	13,844	15,379	17,292	19,820	21,792	25,336	29,246	31,795	35,563	38,885	39,889	41,923	42,856	45,642
27	12,879	14,125	14,573	16,151	18,114	20,703	22,719	26,336	30,319	32,912	36,741	40,113	41,132	43,194	44,140	46,963
28	13,565	14,847	15,308	16,928	18,939	21,588	23,647	27,336	31,319	34,027	37,916	41,337	42,370	44,461	45,419	48,278
29	14,256	15,574	16,047	17,708	19,768	22,475	24,577	28,336	32,461	35,139	39,087	42,557	43,604	45,722	46,693	49,588
30	14,953	16,306	16,791	18,493	20,599	23,364	25,508	29,336	33,530	36,250	40,256	43,773	44,834	46,979	47,962	50,892
n	p=99%	98%	97,5%	95%	90%	80%	70%	50%	30%	20%	10%	5%	4%	2,5%	2%	1%

n = graus de liberdade.

Esta tabela foi adaptada do livro Estatística Básica - Métodos Quantitativos. Bussab, W.O. e Morettin, P.A. Atual Editora.